DOCUMENT RESUME

ED 236 207 TM 830 716

AUTHOR Lord, Frederic M.; Wingersky, Marilyn S.
TITLE Comparison of IRT Observed-Score and True-Score 'Equatings.'
INSTITUTION Educational Testing Service, Princeton, N.J.
SPONS AGENCY Office of Naval Research, Arlington, Va. Personnel and Training Research Programs Office.
REPORT NO ETS-RR-83-26-ONR
PUB DATE Jul 83
CONTRACT N00014-80-C-0402
NOTE 38p.
PUB TYPE Reports - Research/Technical (143)

EDRS PRICE MF01/PC02 Plus Postage.
DESCRIPTORS *Comparative Analysis; *Equated Scores; Estimation (Mathematics); *Latent Trait Theory; Measurement Techniques; Reliability; *Scores; Statistical Analysis; *True Scores

ABSTRACT

Two methods of 'equating' tests using item response theory (IRT) are compared, one using true scores, the other using the estimated distribution of observed scores. On the data studied, they yield almost indistinguishable results. This is a reassuring result for users of IRT equating methods. (Author)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

ED236207

RR-83-26-ONR

COMPARISON OF IRT OBSERVED-SCORE AND TRUE-SCORE 'EQUATINGS'

Frederic M. Lord

and

Marilyn S. Wingersky

This research was sponsored in part by the Personnel and Training Research Programs Psychological Sciences Division Office of Naval Research, under Contract No. N00014-80-C-0402

Contract Authority Identification Number NR No. 150-453

Frederic M. Lord, Principal Investigator

Educational Testing Service

Princeton, New Jersey

July 1983

TM 830 716

U.S. DEPARTMENT OF EDUCATION
NATIONAL INSTITUTE OF EDUCATION
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

X This document has been reproduced as received from the person or organization originating it.

☐ Minor changes have been made to improve reproduction quality.

• Points of view or opinions stated in this document do not necessarily represent official NIE position or policy.

Reproduction in whole or in part is permitted for any purpose of the United States Government.

Approved for public release; distribution unlimited.

Unclassified

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

REPORT DOCUMENTATION PAGE		READ INSTRUCTIONS BEFORE COMPLETING FORM
1. REPORT NUMBER	2. GOVT ACCESSION NO.	3. RECIPIENT'S CATALOG NUMBER
4. TITLE (and Subtitle) Comparison of IRT Observed-Score and True-Score 'Equatings'		5. TYPE OF REPORT & PERIOD COVERED Technical Report 6. PERFORMING ORG. REPORT NUMBER RR-83-26-ONR
7. AUTHOR(s) Frederic M. Lord and Marilyn S. Wingersky		8. CONTRACT OR GRANT NUMBER(s) N00014-80-C-0402
9. PERFORMING ORGANIZATION NAME AND ADDRESS Educational Testing Service Princeton, New Jersey 08541		10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS NR 150-453
11. CONTROLLING OFFICE NAME AND ADDRESS Personnel and Training Research Programs Office of Naval Research Arlington, Virginia 22217		12. REPORT DATE July 1983 13. NUMBER OF PAGES 21
14. MONITORING AGENCY NAME & ADDRESS(if different from Controlling Office)		15. SECURITY CLASS. (of this report) Unclassified 15a. DECLASSIFICATION/DOWNGRADING SCHEDULE

16. DISTRIBUTION STATEMENT (of this Report)

Approved for public release; distribution unlimited.

17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)

18. SUPPLEMENTARY NOTES

19. KEY WORDS (Continue on reverse side if necessary and identify by block number)

Item Response Theory
Equating

20. ABSTRACT (Continue on reverse side if necessary and identify by block number)

Two methods of 'equating' tests using item response theory are compared, one using true scores, the other using estimated observed scores. On the data studied, they yield almost indistinguishable results. This is a reassuring result for users of IRT equating methods.

DD FORM 1 JAN 73 1473 EDITION OF 1 NOV 65 IS OBSOLETE
S/N 0102-LF-014-6601

Unclassified

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

Comparison of IRT Observed-Score and True-Score 'Equatings'

Abstract

Two methods of 'equating' tests using item response theory are compared, one using true scores, the other using the estimated distribution of observed scores. On the data studied, they yield almost indistinguishable results. This is a reassuring result for users of IRT equating methods.

Comparison of IRT Observed-Score and True-Score 'Equatings'*

Most IRT equating is currently attempted by the true-score equating procedure described in Lord (1980, Chapter 13). Lord also describes an IRT observed-score procedure, which until now seems not to have been further investigated, perhaps because it is more complicated and more expensive than the true-score procedure. The present article reports an empirical research study comparing the results of applying these two procedures to real test data.

Sections 1 and 2 outline the true-score and the observed-score procedures, respectively. Section 3 discusses the theoretical advantages and disadvantages of each procedure. Section 4 describes the real test data used to provide a comparison of the two methods. Section 5 describes the procedures for estimating item and ability parameters. Section 6 reports and summarizes the empirical results.

Item response theory models the probability of a correct response by an examinee to a test item as a monotonically increasing function of ability. The model used here is Birnbaum's three-parameter logistic model given by the following formula:

---

*This work was supported in part by contract N00014-80-C-0402, project designation NR 150-453 between the Office of Naval Research and Educational Testing Service. Reproduction in whole or in part in permitted for any purpose of the United States Government.

$$P_i(\theta_a) = c_i + (1 - c_i)/(1 + \exp(-1.7a_i(\theta_a - b_i))) \qquad (1)$$

where $P_i(\theta_a)$ is the probability of examinee a getting item i correct

b_i is the difficulty of item i ;

a_i is the discrimination index for item i ;

c_i is the lower asymptote for item i ;

θ_a is the ability of examinee a ($-\infty < \theta_a < \infty$).

$P_i(\theta_a)$ has a minimum of c_i and a maximum of 1. This model assumes that the test is unidimensional.

1. True-Score Equating

Since the expected score of examinee a on item i is $P_i(\theta_a)$, the examinee's expected number of right answers is $\Sigma_i P_i(\theta_a)$. In classical test theory, this expectation is called the (number-right) true score, $\xi_a \equiv \Sigma_i P_i(\theta_a)$. For the moment, we do not deal with the scores of particular examinees, so the subscript a will be dropped. Here the true score for test X containing n items is the mathematical variable

$$\xi = \sum_{i=1}^{n} P_i(\theta) \quad , \qquad (2)$$

a monotonic increasing function of θ. If test Y contains m items and measures the same ability θ as test X, the true score on test Y is the mathematical variable

$$\eta \equiv \sum_{j=1}^{m} P_j(\theta) \quad . \tag{3}$$

The variables ξ, η, θ are all measures of the same psychological trait, they differ only in the numerical scale on which the measurements are expressed. Thus true scores $\xi = \xi_o$ and $\eta = \eta_o$ corresponding to any given $\theta = \theta_o$ represent identical levels of ability. Any examinee whose true score on test X is ξ_o must automatically have a true score on test Y of exactly η_o, provided the IRT model holds. The situation is the same as when we say that 32° Farenheit has the same meaning as 0° Celsius, except that these temperature scales have a linear relationship, whereas the true-score scales have a nonlinear relationship. Thus, ξ_0 and η_0 are equated true scores; this is true in a much stronger sense than is usually implied by the term equated.

In IRT true-score equating, estimated item parameters are substituted into (2) and (3) and a table of corresponding values of ξ and η is calculated. This constitutes the true-score equating table. This table is then applied in practice as if the true scores

were observed number-right scores. Since observed scores have different properties than true scores, this last step has no clear theoretical justification. It is done as a practical procedure, to be justitied only by whatever usefulness and reasonableness can be empirically demonstrated for the results.

2. IRT Observed-Score Equating

If the assumptions of IRT hold (as is assumed throughout), the probability that an examinee of ability θ will have a number-right score of $x = 1$ on a two-item test is $P_1Q_2 + Q_1P_2$, where $P_i \equiv P_i(\theta)$ and $Q_i \equiv 1 - P_i$. The probability that this examinee's score is 0 or 1 is Q_1Q_2 or P_1P_2 respectively. These probabilities constitute the conditional frequency distribution $f_2(x|\theta)$.

If a third item is added to this test, the distribution of x is now

$$f_3(x|\theta) = Q_3 f_2(x|\theta) + P_3 f_2(x - 1|\theta) \qquad (x = 0,1,\ldots,3) .$$

where $f_r(x|\theta) = 0$ if $x < 0$ or $x > r$. Using this recursive procedure, a computer can readily determine $f_r(x|\theta)$, even for an n of several hundred.

If the θ of each examinee is known, the (marginal) distribution of x for a group of N examinees is

$$\frac{1}{N} \sum_{a=1}^{N} f_n(x|\theta_a) \quad . \tag{4}$$

If an m-item test Y yields number-right score y and measures the same ability as test X, then the (marginal) distribution of y for a group of M examinees is

$$\frac{1}{M} \sum_{b=1}^{M} f_m(y|\theta_b) \quad . \tag{5}$$

A monotonic transformation of the y scores can now be found from (4) and (5) such that the distribution of the transformed y scores is the same as the distribution of the (untransformed) x scores, except for irregularities due to the fact that x and y can only assume integer values. This is done by finding, for each y score, the x that has the same percentile rank in (4) that y has in (5). The x so found is the desired transformed y score.

If the examinees who took test Y have the same distribution of θ as the examinees who took test X, then the resulting transformation of y is an 'equipercentile equating' of the y scale to the x scale. Within groups similar to the groups used to derive the transformation, it has the valuable property that if a cutting score

is chosen on the x scale and the same cutting score is used on the transformed y scale, the proportion of test X examinees selected will be the same as the proportion of test Y examinees selected. This property is essential if test X and test Y examinees are both to be treated equitably, so that an examinee cannot complain that he was injured by the choice of test administered.

When the groups taking tests X and Y are known to have approximately the same distribution of θ (for example, they are two random samples from the same population), there is no reason to use IRT. It is much simpler to do the equipercentile equating using the actual sample distributions of x and y, instead of (4) and (5). The need for IRT arises when the ability distributions of the two groups may differ. In this case, IRT may allow us to estimate the (marginal) frequency distributions of number-right scores that would have resulted if all examinees had taken both tests, without practice or fatigue effects.

In order to do this, the item and ability parameters in (4) and (5) must all be on the same scale. This is usually accomplished by administering a suitable 'anchor test' to both groups of examinees. All answer-sheet responses for both groups are used in a single computer run that estimates all parameters on the same scale. These estimates are then used in (4) and (5), substituting $N + M$ for N or M, to obtain the distributions of x and y for the combined group of $N + M$ examinees. Equipercentile equating of y to x is then carried out in the usual way.

3. Theoretical Perspectives

Practical workers, with the need for equating scores on two different test forms, have over the years used widely different methods (see Angoff, 1971) in an attempt to approximate the desired result. Each practical worker, needing a word to describe his results, asserts that he has produced an equating of y to x. Yet different methods and different groups do not produce identical 'equatings'.

Braun and Holland (1982, page 14) state: "There is some disagreement over what test equating is and the proper method for doing it." They then adopt the definition "Form-X and Form-Y are equated on [population] P " if the distribution of the transformed y scores in population P is the same as the distribution of the (untransformed) x scores.

This definition of the phrase 'equated on population P ' is beyond reproach. One problem, however, is that the qualifying phrase 'on population P ' is typically dropped by the practical worker who writes a research report or publishes an equating table in a test manual.

Unfortunately (as will be shown later in this section) two tests that are equated on population P will typically not be equated for various subpopulations that are included in P . Test scores that are equated for the population of college applicants may well be equated neither for the population of female college applicants, nor for the

the population of male college applicants. The scores are still less likely to be equated for a subpopulation characterized by interest in science, or in music. For the subpopulation of Harvard applicants, the situation is much worse.

If the proportion of applicants admitted to Harvard differs significantly depending on whether they were given form X or form Y of the test, it is clear that the 'equating' was unsuccessful. Since similar inequalities are likely to characterize any equating on any specified population, it may be best not to say that the tests are 'equated' at all, or to simply say that they are 'approximately equated.'

From a practical point of view, the approximation may be quite satisfactory for many subgroups. It is unlikely, however, that the equating will be adequate for any subpopulation having a mean and variance of ability that is sharply different from the mean and variance of the total population used to derive the equating transformation. Extensive practical data illustrating the adequacies and the inadequacies of approximate equatings are given in the 30-volume Anchor Test Study (Loret, Seder, Bianchini, and Vale, 1974).

For a theoretical discussion of alternative equating methods, however, it is important not to start out with a definition of equating that is clearly inadequate for subpopulations of examinees. Given that the IRT model holds, IRT observed-score equating would, for example, be automatically endorsed by the Braun and Holland definition, since their definition mandates equipercentile equating. IRT true-score equating would be definitely rejected by their definition, since in general it

will not lead to x scores and transformed y scores having the same frequency distribution, unless X and Y are strictly parallel forms that are identical in difficulty, in reliability, and also in most other respects.

The important virtue of IRT true-score equating is that if the IRT model holds, the true scores are clearly equated for all subpopulations of examinees. This results from the invariance of IRT parameters across populations of examinees, assumed by the IRT model. The clear flaw in IRT true-score equating is that it equates true scores, not the actually observed fallible scores. Treating observed scores as if they were true scores cannot be justified on any theoretical grounds.

The virtue of IRT observed-score equating is that in a group like that used to derive the equating, any cutting score will accept the same percentage of examinees regardless of the test administered. The flaw is that this holds only for that total group and not for other groups or subgroups.

This last statement is most clearly seen from a very extreme example. Suppose forms X and Y have the same number of items, measure the same ability θ, but differ in difficulty. If the equipercentile equating is carried out on a group of examinees all of whom are guessing at random on almost all the items, the difference in difficulty between the two forms will not manifest itself and any equipercentile equating will approximate an identity transformation of score y. If a slightly more competent group of examinees is used for the equipercentile equating, however, the difference in difficulty

between forms will begin to become apparent and most y scores will be adjusted upwards or downwards accordingly. As the competence of the group used becomes higher and higher, the equating transformation found will differ more and more from the identity transformation found from the original extreme group.

As a second example of the inescapable invalidity of observed-score equating, suppose that tests X and Y are of equal difficulty and that the true scores ξ and η have equal variance, but that y is much less reliable than x. Consider a subgroup of very talented examinees; to make the illustration clear, consider that in this subgroup all examinees have nearly identical θ values. Most of the variation in observed scores x and y is now due to errors of measurement. The equipercentile equating transformation found will thus approximate a straight line with slope

$$\frac{\text{standard deviation of the errors of measurement in } x}{\text{standard deviation of the errors of measurement in } y}.$$

Since y is much less reliable than x, the slope will be much less than 1.

If, on the other hand, the equipercentile equating transformation is found from a group where the true-score variance is large compared to the errors variances, the transformation will tend to approximate a straight line with slope

$$\frac{\text{standard deviation of true scores on } x}{\text{standard deviation of true scores on } y}.$$

Intermediate situations will provide transformations with intermediate slopes. If the wrong equating is applied to any given subpopulation, the population of examinees in the subpopulation accepted will depend on whether they took test X or test Y , an inequitable result.

Our theoretical position, then, is that each method described in Section 2 (as well as all other available equating methods) has its own inadequacies. Since, in practice, some (approximate) equating method must be used, it will be informative to investigate empirically how the two methods of Section 2 compare in a specially contrived practical situation where the correct equating is actually known in advance.

4. Data

These two equating methods were used to equate the chain of six SAT verbal tests described by Petersen, Cook and Stocking in the report IRT Versus Conventional Equating Methods: A Comparative Study of Scale Stability. The tests in this chain were selected such that the first test and the last test are the same. Each test is equated to the next test in the chain using an anchor test. Figure 1 is a diagram of the chain. The capital letters represent the test form, the small letters represent the anchor test. Scores on form V4 are equated to scores on form X2 using the anchor test fe . These equated scores on X2 are equated to scores on form Y3 using the anchor test fm . This gives us an equating of form V4 to Y3 . In this manner, one proceeds through the chain, with the final equating of Z5 to V4 giving us a

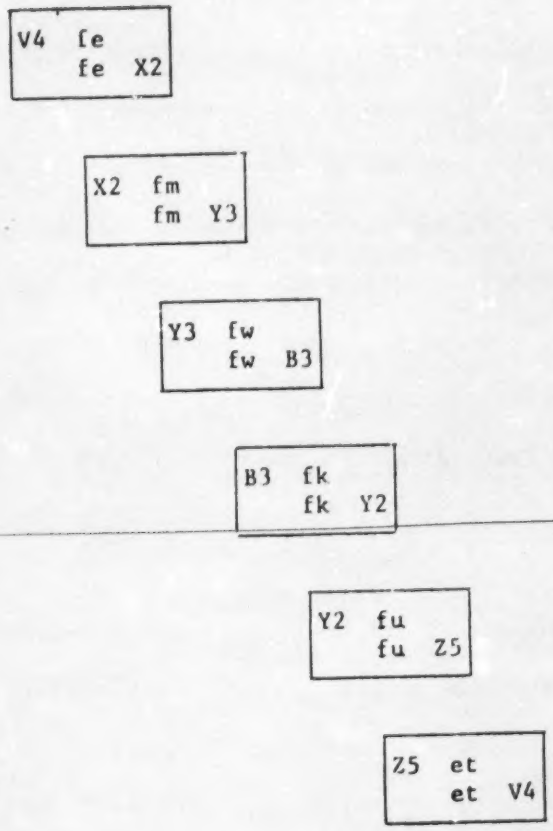


Figure 1. Chain of six SAT verbal equatings. Upper case letters designate test forms; lower case letters designate anchor tests.

table of scores on the original V4 equated to the scores on the V4 at the end of the chain. Any deviation from equality between the two sets of scores could be attributable to scale drift or lack of model fit.

Each form in the chain has 85 items except form V4 which has 90 items. Each anchor test has 40 items. For each form there are two samples of examinees; each sample taking a different anchor test. The two groups taking each form were random samples from the same population for all of the forms except Y3 . For the parameter estimation runs a random sample of approximately 2670 examinees was selected from the data obtained at the test administration of that form and anchor test.

5. Parameter Calibration

The item parameters and abilities were estimated by a modified version of the computer program LOGIST, (Wood, Wingersky, & Lord, 1976) in six separate calibration runs. In Figure 1, each box (containing two forms and one anchor test) represents one LOGIST run. The item responses for items not taken by an examinee, such as the X2 items for examinees taking form V4 in box 1, are treated as not reached items.

All of the estimated parameters within each LOGIST run are on the same scale and either method of equating can be used to equate the scores for the two tests. The anchor tests are not used directly in the equating but are used in LOGIST so that the estimated parameters within a LOGIST run are on the same scale.

6. Results

In using the IRT observed-score equating method, two estimated distributions of observed scores are equated so that the transformed y scores and the (untransformed) x scores have the same distribution. Figure 2 is presented to demonstrate, that at least for one set of data, this estimated distribution of observed scores is a reasonable fit to the actual distribution of observed scores. The frequencies are plotted against formula scores which are the number right minus a fraction of the number wrong. The fraction is one over the number of choices. Since the estimated observed-score distribution can only be obtained for number-right scores, the transformation to formula scores assumes that there are no omits, that is, that the number wrong is the total number of items minus the number right. In order to compare the two distributions, the observed-score distribution should be based on a group that has no omits. Consequently, a form of the SAT verbal different from the ones in the chain was used for this Figure in order to get a sufficiently large enough sample for the frequency distribution and for the item calibration.

The agreement shown in Figure 2 is good except that the tails of the estimated distribution are too high. This discrepancy is presumably due to the use of estimated θ in place of true θ for the practical implementation of (4). Since a similar discrepancy affects

FIGURE 2

COMPARISON OF DISTRIBUTION OF OBSERVED SCORES AND ESTIMATED DISTRIBUTION OF OBSERVED SCORES - SAT VERBAL - NO OMITS

---- ESTIMATED OBSERVED-SCORE DISTRIBUTION
—— OBSERVED-SCORE DISTRIBUTION

the estimated observed-score distributions of both test X and test Y , the effects of the discrepancies tend to cancel out in the equating process.

In our chain-equating study, each method of equating was applied separately to the whole chain of equatings, resulting in a line for each method equating form V4 at the beginning of the chain to form V4 at the end of the chain. These two lines are plotted in Figure 3 along with a 45° line. The solid line is the IRT true-score equating line; the dotted line, falling practically on top of the solid line, is the IRT observed-score equating line. To equate scores below chance level, that is $\sum_{i=1}^{n} c_i$, for the IRT true-score line, the method given on pages 210-211 of Lord (1980) was used. For scores above 0, the maximum difference between the two equatings was .2; for scores below 0, the maximum difference was .8 which occurred at the chance level. If the equating methods were perfect and there were no scale drift, the equating line would be the dashed 45° line.

Figure 4 shows the two equating methods applied to one individual link in the chain. This particular link was selected because the IRT true-score equating line between these two forms had the greatest discontinuity in the slope at the chance level. The largest difference between the two lines occurred at the chance level and was 1.6. For scores above 0 the maximum difference between the two lines was .4.

FIGURE 3
COMPARISON OF TRUE-SCORE EQUATING
AND ESTIMATED OBSERVED-SCORE EQUATING
OVER CHAIN OF SIX EQUATINGS

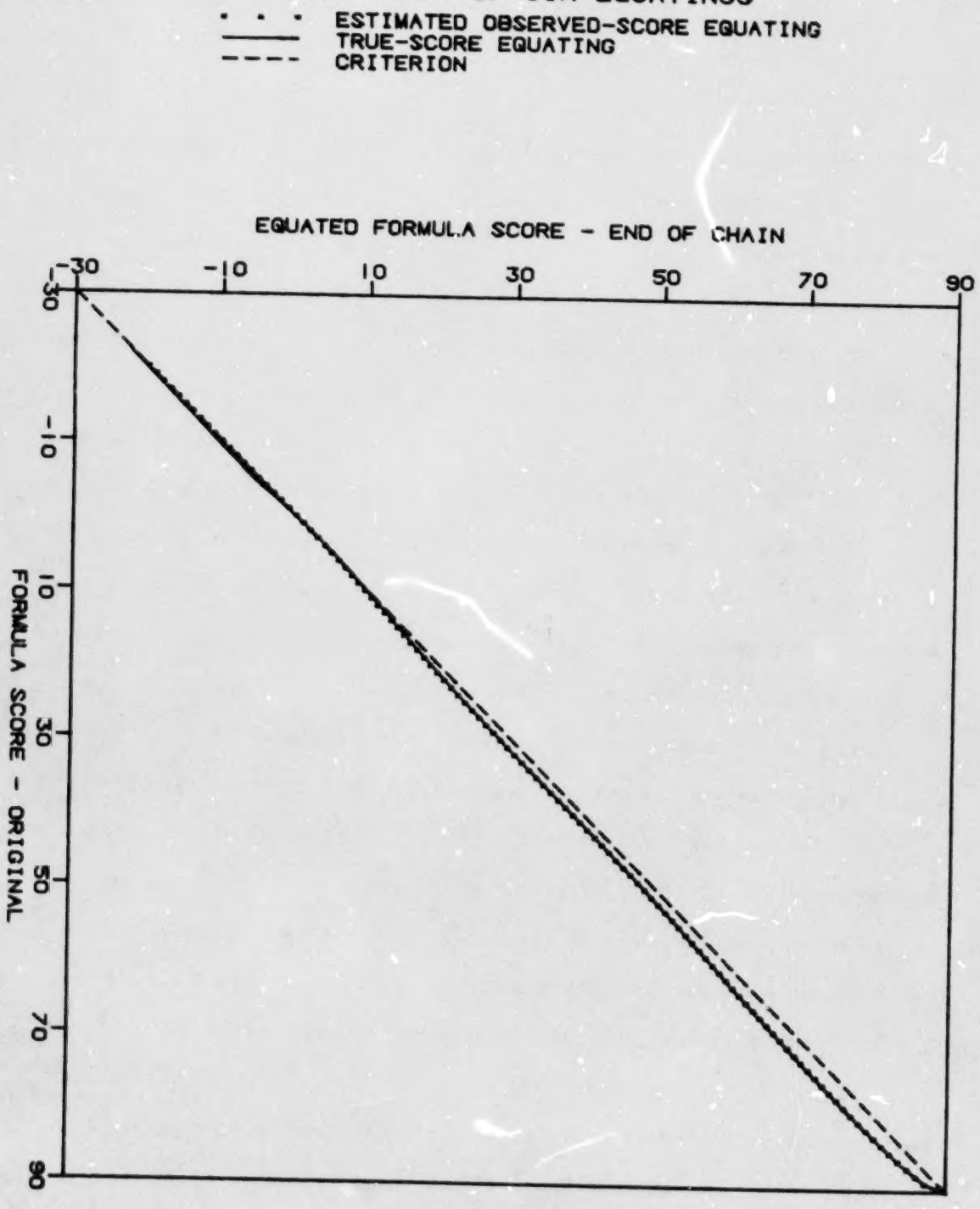
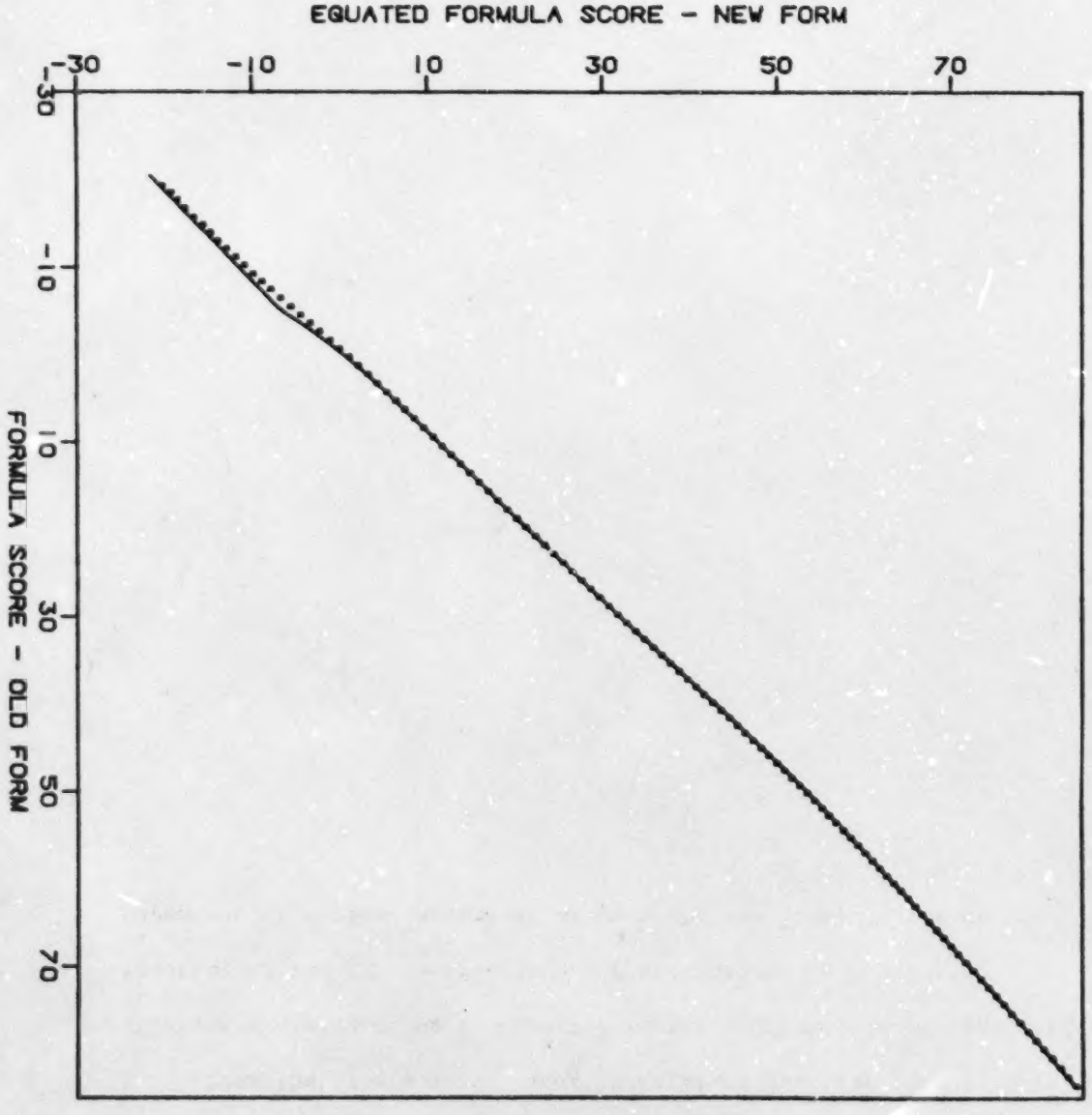


FIGURE 4
COMPARISON OF TRUE-SCORE EQUATING
AND ESTIMATED OBSERVED-SCORE EQUATING
SINGLE LINK IN CHAIN

• • • ESTIMATED OBSERVED-SCORE EQUATING
——— TRUE-SCORE EQUATING



Given that there is no clear theoretical justification for applying IRT true-score equating to observed scores and that the equipercentile equating of the IRT observed-score distributions is population dependent, the close agreement between the two lines is reassuring.

References

Angoff, W. H. Scales, norms, and equivalent scores. In R. L. Thorndike (Ed.), Educational measurement (2nd ed.). Washington, D.C.: American Council on Education, 1971. Pp. 508-600.

Braun, H. I. & Holland, P. W. Observed-score test equating: A mathematical analysis of some ETS equating procedures. In P. W. Holland and D. B. Rubin (Eds.), Test equating. New York: Academic Press, 1982. Pp. 9-49.

Lord, F. M. Applications of item response theory to practical testing problems. Hillsdale, N.J.: Lawrence Erlbaum Associates, 1980.

Loret, P. G., Seder, A., Bianchini, J. C., & Vale, C. A. Anchor test study--Equivalence and norms tables for selected reading achievement tests (grades 4, 5, 6). Washington, D.C.: United States Government Printing Office, 1974.

Petersen, N. S., Cook, L. L., & Stocking, M. L. IRT versus conventional equating methods: A comparative study of scale stability. Paper presented at the annual meeting of the American Educational Research Association, April 1982.

Wood, R L, Wingersky, M. S., & Lord, F. M. LOGIST--A computer program for estimating examinee ability and item characteristic curve parameters. (ETS RM 76-6). Princeton, N.J.: Educational Testing Service, 1976.

DISTRIBUTION LIST

Navy

1 Dr. Ed Aiken
Navy Personnel R&D Center
San Diego, CA 92152

1 Dr. Arthur Bachrach
Environmental Stress Program Center
Naval Medical Research Institute
Bethesda, MD 20014

1 Dr. Meryl S. Baker
Navy Personnel R&D Center
San Diego, CA 92152

1 Liaison Scientist
Office of Naval Research
Branch Office London
Box 39
FPO New York, NY 09510

1 Lt. Alexander Bory
Applied Psychology
Measurement Division
NAMRL
NAS Pensacola, FL 32508

1 Dr. Robert Breaux
NAVTRAEQUIPCEN
Code N-095R
Orlando, FL 32813

1 Dr. Robert Carroll
NAVOP 115
Washington, DC 20370

1 Chief of Naval Education and
Training Liason Office
Air Force Human Resource Laboratory
Flying Training Division
Williams Air Force Base, AZ 85224

1 Dr. Stanley Collyer
Office of Naval Technology
800 N. Quincy Street
Arlington, VA 22217

1 CDR Mike Curran
Office of Naval Research
800 North Quincy Street
Code 270
Arlington, VA 22217

1 Dr. Tom Duffy
Navy Personnel R&D Center
San Diego, CA 92152

1 Mr. Mike Durmeyer
Instructional Program Development
Building 90
NET-PDCD
Great Lakes NTC, IL 60088

1 Dr. Richard Elster
Department of Administrative Sciences
Naval Postgraduate School
Monterey, CA 93940

1 Dr. Pat Federico
Code P13
Navy Personnel R & D Center
San Diego, CA 92152

1 Dr. Cathy Fernandes
Navy Personnel R & D Center
San Diego, CA 92152

1 Dr. John Ford
Navy Personnel R & D Center
San Diego, CA 92152

1 Dr. Jim Hollan
Code 14
Navy Personnel R & D Center
San Diego, CA 92152

1 Dr. Ed Hutchins
Navy Personnel R & D Center
San Diego, CA 92152

1 Dr. Norman J. Kerr
Chief of Naval Technical Training
Naval Air Station Memphis (75)
Millington, TN 38054

1 Dr. Peter Kincaid
Training Analysis & Evaluation Group
Department of the Navy
Orlando, FL 32813

1 Dr. R. W. King
Director, Naval Education
and Training Program
Naval Training Center, Bldg. 90
Great Lakes, IL 60088

1 Dr. Leonard Kroeker
Navy Personnel R & D Center
San Diego, CA 92152

1 Dr. William L. Maloy (02)
Chief of Naval Education and Training
Naval Air Station
Pensacola, FL 32508

1 Dr. Kneale Marshall
Chairman, Operations Research Dept.
Naval Post Graduate School
Monterey, CA 93940

1 Dr. James McBride
Navy Personnel R & D Center
San Diego, CA 92152

1 Dr. William Montague
NPRDC Code 13
San Diego, CA 92152

1 Mr. William Nordbrock
1032 Fairlawn Avenue
Libertyville, IL 60048

1 Library, Code P201L
Navy Personnel R & D Center
San Diego, CA 92152

1 Technical Director
Navy Personnel R & D Center
San Diego, CA 92152

6 Personnel & Training Research Group
Code 442PT
Office of Naval Research
Arlington, VA 22217

1 Special Asst. for Education and
Training (OP-01E)
Room 2705 Arlington Annex
Washington, DC 20370

1 LT Frank C. Petho, MSC, USN
CNET (N-432)
NAS
Pensacola, FL 32508

1 Dr. Bernard Rimland (01C)
Navy Personnel R & D Center
San Diego, CA 92152

1 Dr. Carl Ross
CNET-PDCD
Building 90
Great Lakes NTC, IL 60088

1 Dr. Worth Scanland, Director
CNET (N-5)
NAS
Pensacola, FL 32508

1 Dr. Robert G. Smith
Office of Chief of Naval Operations
OP-987H
Washington, DC 20350

1 Dr. Alfred F. Smode, Director
Training Analysis and Evaluation Group
Department of the Navy
Orlando, FL 32813

1 Dr. Richard Sorensen
Navy Personnel R & D Center
San Diego, CA 92152

1 Dr. Frederick Steinheiser
CNO - OP115
Navy Annex
Arlington, VA 20370

1 Mr. Brad Sympson
Naval Personnel R & D Center
San Diego, CA 92152

1 Dr. Frank Vicino
Navy Personnel R & D Center
San Diego, CA 92152

1 Dr. Edward Wegman
Office of Naval Research (Code 411S&P)
800 North Quincy Street
Arlington, VA 22217

1 Dr. Ronald Weitzman
Code 54 WZ
Department of Administrative Services
U.S. Naval Postgraduate School
Monterey, CA 93940

1 Dr. Douglas Wetzel
Code 12
Navy Personnel R & D Center
San Diego, CA 92152

1 Dr. Martin F. Wiskoff
Navy Personnel R & D Center
San Diego, CA 92152

1 Mr. John H. Wolfe
Navy Personnel R & D Center
San Diego, CA 92152

1 Dr. Wallace Wulfeck, III
Navy Personnel R & D Center
San Diego, CA 92152

Marine Corps

1 Dr. H. William Greenup
Education Advisor (E031)
Education Center, MCDEC
Quantico, VA 22134

1 Director, Office of Manpower Utilization
HQ, Marine Corps (MPU)
BCB, Building 2009
Quantico, VA 22134

1 Headquarters, U. S. Marine Corps
Code MPI-20
Washington, DC 20380

1 Special Assistant for Marine Corps Matters
Code 100M
Office of Naval Research
800 N. Quincy Street
Arlington, VA 22217

1 Dr. A. L. Slafkosky
Scientific Advisor
Code RD-1
HQ, U.S. Marine Corps
Washington, DC 20380

1 Major Frank Yohannan, USMC
Headquarters, Marine Corps
(Code MPI-20)
Washington, DC 20380

Army

1 Technical Director
U.S. Army Research Institute for the Behavioral and Social Sciences
5001 Eisenhower Avenue
Alexandria, VA 22333

1 Mr. James Baker
U.S. Army Research Institute
5001 Eisenhower Avenue
Alexandria, VA 22333

1 Dr. Kent Eaton
U.S. Army Research Institute
5001 Eisenhower Avenue
Alexandria, VA 22333

Army

1 Dr. Beatrice J. Farr
U.S. Army Research Institute
5001 Eisenhower Avenue
Alexandria, VA 22333

1 Dr. Myron Fischl
U.S. Army Research Institute for the Social and Behavioral Sciences
5001 Eisenhower Avenue
Alexandria, VA 22333

1 Dr. Milton S. Katz
Training Technical Area
U.S. Army Research Institute
5001 Eisenhower Avenue
Alexandria, VA 22333

1 Dr. Harold F. O'Neil, Jr.
Director, Training Research Lab
U.S. Army Research Institute
5001 Eisenhower Avenue
Alexandria, VA 22333

1 Commander, U.S. Army Research Institute
ATTN: PERI-BR (Dr. Judith Orasanu)
5001 Eisenhower Avenue
Alexandria, VA 22333

1 Dr. Joseph Psotka
ATTN: PERI-1C
U.S. Army Research Institute
5001 Eisenhower Avenue
Alexandria, VA 22333

1 Mr. Robert Ross
U.S. Army Research Institute for the Social and Behavioral Sciences
5001 Eisenhower Avenue
Alexandria, VA 22333

1 Dr. Robert Sasmor
U.S. Army Research Institute for the Social and Behavioral Sciences
5001 Eisenhower Avenue
Alexandria, VA 22333

1 Dr. Joyce Shields
U.S. Army Research Institute
5001 Eisenhower Avenue
Alexandria, VA 22333

1 Dr. Hilda Wing
U.S. Army Research Institute
5001 Eisenhower Avenue
Alexandria, VA 22333

1 Dr. Robert Wisher
U.S. Army Research Institute
5001 Eisenhower Avenue
Alexandria, VA 22333

Air Force

1 Air Force Human Resources Laboratory
AFHRL/MPD
Brooks Air Force Base, TX 78235

1 Technical Documents Center
Air Force Human Resources Laboratory
WPAFB, OH 45433

1 U.S. Air Force Office of Scientific Research
Life Sciences Directorate, NL
Bolling Air Force Base
Washington, DC 20332

1 Air University Library
AUL/LSE 76/443
Maxwell AFB, AL 36112

1 Dr. Earl A. Alluisi
HQ, AFHRL (AFSC)
Brooks Air Force Base, TX 78235

1 Mr. Raymond E. Christal
AFHRL/MOE
Brooks AFB, TX 78235

1 Dr. Alfred R. Fregly
AFOSR/NL
Bolling AFB, DC 20332

Air Force

1 Dr. Genevieve Haddad
Program Manager
Life Sciences Directorate
AFOSR
Bolling AFB, DC 20332

1 Dr. T. M. Longridge
AFHRL/OTE
Williams AFB, AZ 85224

1 Dr. Roger Pennell
Air Force Human Resources Laboratory
Lowry AFB, CO 80230

1 Dr. Malcolm Ree
AFHRL/MP
Brooks Air Force Base, TX 78235

1 LT Tallarigo
3700 TCHTW/TTGHR
Sheppard AFB, TX 76311

1 Dr. Joseph Yasatuke
AFHRL/LRT
Lowry AFB, CO 80230

Department of Defense

12 Defense Technical Information Center
Attn: TC
Cameron Station, Building 5
Alexandria, VA 22314

1 Dr. Craig I. Fields
Advanced Research Projects Agency
1400 Wilson Blvd.
Arlington, VA 22209

1 Dr. William Graham
Testing Directorate
MEPCOM/MEPCT-P
Ft. Sheridan, IL 60037

Department of Defense

1 Mr. Jerry Lehnus
HQ MEPCOM
Attn: MEPCT-P
Ft. Sheridan, IL 60037

1 Military Assistant for Training
and Personnel Technology
Office of the Under Secretary of
Defense for Research and Engineering
Room 3D129, The Pentagon
Washington, DC 20301

1 Dr. Wayne Sellman
Office of the Assistant Secretary
of Defense (MRA&L)
2B269 The Pentagon
Washington, DC 20301

1 Major Jack Thorpe
DARPA
1400 Wilson Blvd.
Arlington, VA 22209

Civilian Agencies

1 Dr. Patricia A. Butler
NIE-BRN Bldg., Stop #7
1200 19th Street, NW
Washington, DC 20208

1 Dr. Susan Chipman
Learning and Development
National Institute of Education
1200 19th Street NW
Washington, DC 20208

1 Dr. Arthur Melmed
724 Brown
U.S. Department of Education
Washington, DC 20208

1 Dr. Andrew R. Molnar
Office of Scientific and Engineering
Personnel and Education
National Science Foundation
Washington, DC 20550

Civilian Agencies

1 Dr. Vern W. Urry
Personnel R & D Center
Office of Personnel Management
1900 E Street, NW
Washington, DC 20415

1 Mr. Thomas A. Warm
U.S. Coast Guard Institute
P.O. Substation 18
Oklahoma City, OK 73169

1 Dr. Frank Withrow
U.S. Office of Education
400 Maryland Avenue, SW
Washington, DC 20202

1 Dr. Joseph L. Young, Director
Memory and Cognitive Processes
National Science Foundation
Washington, DC 20550

Private Sector

1 Dr. James Algina
University of Florida
Gainesville, FL 32611

1 Dr. Patricia Baggett
Department of Psychology
University of Colorado
Boulder, CO 80309

1 Dr. Isaac Bejar
Educational Testing Service
Princeton, NJ 08541

1 Dr. Menucha Birenbaum
School of Education
Tel Aviv University
Tel Aviv, Ramat Aviv 69978
ISRAEL

Private Sector

1 Dr. R. Darrell Bock
Department of Education
University of Chicago
Chicago, IL 60637

1 Dr. Robert Brennan
American College Testing Programs
P.O. Box 168
Iowa City, IA 52243

1 Dr. Glenn Bryan
6208 Poe Road
Bethesda, MD 20817

1 Dr. Ernest R. Cadotte
307 Stokely
University of Tennessee
Knoxville, TN 37916

1 Dr. Pat Carpenter
Department of Psychology
Carnegie-Mellon University
Pittsburgh, PA 15213

1 Dr. John B. Carroll
409 Elliott Road
Chapel Hill, NC 27514

1 Dr. Norman Cliff
Department of Psychology
University of Southern California
University Park
Los Angeles, CA 90007

1 Dr. Allan M. Collins
Bolt, Beranek, and Newman, Inc.
50 Moulton Street
Cambridge, MA 02138

1 Dr. Lynn A. Cooper
LRDC
University of Pittsburgh
3939 O'Hara Street
Pittsburgh, PA 15213

Private Sector

1 Dr. Hans Crombag
Education Research Center
University of Leyden
Boerhaavelaan 2
2334 EN Leyden
THE NETHERLANDS

1 Dr. Dattpradad Divgi
Syracuse University
Department of Psychology
Syracuse, NY 33210

1 Dr. Susan Embertson
Psychology Department
University of Kansas
Lawrence, KS 66045

1 ERIC Facility-Acquisitions
4833 Rugby Avenue
Bethesda, MD 20014

1 Dr. Benjamin A. Fairbank, Jr.
McFann-Gray and Associates, Inc.
5825 Callaghan
Suite 225
San Antonio, TX 78228

1 Dr. Leonard Feldt
Lindquist Center for Measurement
University of Iowa
Iowa City, IA 52242

1 Prof. Donald Fitzgerald
University of New England
Armidale, New South Wales 2351
AUSTRALIA

1 Dr. Dexter Fletcher
WICAT Research Institute
1875 S. State Street
Orem, UT 22333

1 Dr. John R. Frederiksen
Bolt, Beranek, and Newman
50 Moulton Street
Cambridge, MA 02138

Private Sector

1 Dr. Janice Gifford
University of Massachusetts
School of Education
Amherst, MA 01002

1 Dr. Robert Glaser
LRDC
University of Pittsburgh
3939 O'Hara Street
Pittsburgh, PA 15213

1 Dr. Bert Green
Department of Psychology
Johns Hopkins University
Charles and 34th Streets
Baltimore, MD 21218

1 Dr. Ron Hambleton
School of Education
University of Massachusetts
Amherst, MA 01002

1 Dr. Paul Horst
677 G Street, #184
Chula Vista, CA 90010

1 Dr. Lloyd Humphreys
Department of Psychology
University of Illinois
Champaign, IL 61820

1 Dr. Jack Hunter
2122 Coolidge Street
Lansing, MI 48906

1 Dr. Huynh Huynh
College of Education
University of South Carolina
Columbia, SC 29208

1 Dr. Douglas H. Jones
10 Trafalgar Court
Lawrenceville, NJ 08648

Private Sector

1 Prof. John A. Keats
Department of Psychology
University of Newcastle
Newcastle, New South Wales 2308
AUSTRALIA

1 Dr. William Koch
University of Texas-Austin
Measurement and Evaluation Center
Austin, TX 78703

1 Dr. Pat Langley
The Robotics Institute
Carnegie-Mellon University
Pittsburgh, PA 15213

1 Dr. Alan Lesgold
Learning R & D Center
University of Pittsburgh
3939 O'Hara Street
Pittsburgh, PA 15260

1 Dr. Michael Levine
Department of Educational Psychology
210 Education Building
University of Illinois
Champaign, IL 61801

1 Dr. Charles Lewis
Faculteit Sociale Wetenschappen
Rijksuniversiteit Groningen
Oude Boteringestraat 23
9712GC Groningen
NETHERLANDS

1 Dr. Robert Linn
College of Education
University of Illinois
Urbana, IL 61801

1 Mr. Phillip Livingston
Systems and Applied Sciences Corporation
68111 Kenilworth Avenue
Riverdale, MD 20840

Private Sector

1 Dr. Robert Lockman
Center for Naval Analysis
200 North Beauregard Street
Alexandria, VA 22311

1 Dr. Frederic M. Lord
Educational Testing Service
Princeton, NJ 08541

1 Dr. James Lumsden
Department of Psychology
University of Western Australia
Nedlands, Western Australia 6009
AUSTRALIA

1 Dr. Gary Marco
Stop 31-E
Educational Testing Service
Princeton, NJ 08541

1 Dr. Scott Maxwell
Department of Psychology
University of Notre Dame
Notre Dame, IN 46556

1 Dr. Samuel T. Mayo
Loyola University of Chicago
820 North Michigan Avenue
Chicago, IL 60611

1 Mr. Robert McKinley
American College Testing Programs
P.O. Box 168
Iowa City, IA 52243

1 Dr. Robert Mislevy
711 Illinois Street
Geneva, IL 60134

1 Dr. Allen Munro
Behavioral Technology Laboratories
1845 Elena Avenue, Fourth Floor
Redondo Beach, CA 90277

Private Sector

1 Dr. Alan Nicewander
University of Oklahoma
Department of Psychology
Oklahoma City, OK 73069

1 Dr. Donald A. Norman
Cognitive Science, C-015
University of California, San Diego
La Jolla, CA 92093

1 Dr. Melvin R. Novick
356 Lindquist Center for Measurement
University of Iowa
Iowa City, IA 52242

1 Dr. James Olson
WICAT, Inc.
1875 S. State Street
Orem, UT 84057

1 Dr. Wayne M. Patience
American Council on Education
GED Testing Service, Suite 20
One Dupont Circle, NW
Washington, DC 20036

1 Dr. James A. Paulson
Portland State University
P.O. Box 751
Portland, OR 97207

1 Dr. James W. Pellegrino
Univeristy of California,
Santa Barbara
Department of Psychology
Santa Barbara, CA 93106

1 Dr. Mark D. Reckase
ACT
P.O. Box 168
Iowa City, IA 52243

1 Dr. Lauren Resnick
LRDC
University of Pittsburgh
3939 O'Hara Street
Pittsburgh, PA 15261

Private Sector

1 Dr. Thomas Reynolds
University of Texas, Dallas
Marketing Department
P.O. Box 688
Richardson, TX 75080

1 Dr. Andrew Rose
American Institutes for Research
1055 Thomas Jefferson St., NW
Washington, DC 20007

1 Dr. Ernst Z. Rothkopf
Bell Laboratories
Murray Hill, NJ 07974

1 Dr. Lawrence Rudner
403 Elm Avenue
Takoma Park, MD 20012

1 Dr. J. Ryan
Department of Education
University of South Carolina
Columbia, SC 29208

1 Prof. Fumiko Samejima
Department of Psychology
University of Tennessee
Knoxville, TN 37916

1 Dr. Walter Schneider
Psychology Department
603 E. Daniel
Champaign, IL 61820

1 Dr. Lowell Schoer
Psychological and Quantitative
Foundations
College of Education
University of Iowa
Iowa City, IA 52242

1 Dr. Robert J. Seidel
Instructional Technology Group
HUMRRO
300 N. Washington Street
Alexandria, VA 22314

Private Sector

1 Dr. Kazuo Shigemasu
University of Tohoku
Department of Educational Psychology
Kawauchi, Sendai 980
JAPAN

1 Dr. Edwin Shirkey
Department of Psychology
University of Central Florida
Orlando, FL 32816

1 Dr. William Sims
Center for Naval Analysis
200 North Beauregard Street
Alexandria, VA 22311

1 Dr. H. Wallace Sinaiko
Program Director
Manpower Research and Advisory Services
Smithsonian Institution
801 North Pitt Street
Alexandria, VA 22314

1 Dr. Richard Snow
School of Education
Stanford University
Stanford, CA 94305

1 Dr. Kathryn T. Spoehr
Psychology Department
Brown University
Providence, RI 02912

1 Dr. Robert Sternberg
Department of Psychology
Yale University
Box 11A, Yale Station
New Haven, CT 06520

1 Dr. Peter Stoloff
Center for Naval Analysis
200 North Beauregard Street
Alexandria, VA 22311

Private Sector

1 Dr. William Stout
University of Illinois
Department of Mathematics
Urbana, IL 61801

1 Dr. Patrick Suppes
Institute for Mathematical Studies
in the Social Sciences
Stanford University
Stanford, CA 94305

1 Dr. Hariharan Swaminathan
Laboratory of Psychometric and
Evaluation Research
School of Education
University of Massacuusetts
Amherst, MA 01003

1 Dr. Kikumi Tatsuoka
Computer Based Education Research
Laboratory
252 Engineering Research Laboratory
University of Illinois
Urbana, IL 61801

1 Dr. Maurice Tatsuoka
220 Education Building
1310 S. Sixth Street
Champaign, Il 61820

1 Dr. David Thissen
Department of Psychology
University of Kansas
Lawrence, KS 66044

1 Dr. Douglas Towne
University of Southern California
Behavioral Technology Labs
1845 S. Elena Avenue
Redondo Beach, CA 90277

1 Dr. Robert Tsutakawa
Department of Statistics
University of Missouri
Columbia, MO 65201

Private Sector

1 Dr. V. R. R. Uppuluri
Union Carbide Corporation
Nuclear Division
P.O. Box Y
Oak Ridge, TN 37830

1 Dr. David Vale
Assessment Systems Corporation
2233 University Avenue
Suite 310
St. Paul, MN 55114

1 Dr. Kurt Van Lehn
Xerox PARC
3333 Coyote Hill Road
Palo Alto, CA 94304

1 Dr. Howard Wainer
Educational Testing Service
Princeton, NJ 08541

1 Dr. Michael T. Waller
Department of Educational Psychology
University of Wisconsin
Milwaukee, WI 53201

1 Dr. Brian Waters
HUMRRO
300 North Washington
Alexandria, VA 22314

1 Dr. Phyllis Weaver
2979 Alexis Drive
Palo Alto, CA 94304

1 Dr. David J. Weiss
N660 Elliott Hall
University of Minnesota
75 East River Road
Minneapolis, MN 55455

1 Dr. Keith T. Wescourt
Perceptronics, Inc.
545 Middlefield Road
Suite 140
Menlo Park, CA 94025

Private Sector

1 Dr. Rand R. Wilcox
University of Southern California
Department of Psychology
Los Angeles, CA 90007

1 Dr. Wolfgang Wildgrube
Streitkraefteamt
Box 20 50 03
D-5300 Bonn 2
WEST GERMANY

1 Dr. Bruce Williams
Department of Educational Psychology
University of Illinois
Urbana, IL 61801

1 Dr. Wendy Yen
CTB/McGraw-Hill
Del Monte Research Park
Monterey, CA 93940